DOCUMENT RESUME

ED 456 157 TM 033 220

AUTHOR	Benford, Russell; Lawson, Anton E.
TITLE	Relationships between Effective Inquiry Use and the Development of Scientific Reasoning Skills in College Biology Labs.
SPONS AGENCY	National Science Foundation, Arlington, VA.
PUB DATE	2001-04-00
NOTE	72p.
CONTRACT	DUE-9453610
PUB TYPE	Reports - Research (143) -- Tests/Questionnaires (160)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Biology; College Faculty; *College Students; Higher Education; *Inquiry; *Teaching Methods; *Thinking Skills
IDENTIFIERS	*Scientific Thinking

ABSTRACT

Two hypotheses regarding the relationship between scientific reasoning skills and the use of the inquiry method of instruction in college biology labs were examined. The first hypothesis was that scientific reasoning skills influence an instructor's ability to teach biology using inquiry. The second hypothesis was that the effectiveness with which an instructor uses inquiry affects the pedagogical outcome of a lesson. To test the first hypothesis, 9 instructors teaching 702 students in an introductory biology course for nonmajors were evaluated for their scientific reasoning skills and understanding of the nature of science. Data were also collected on instructors' prior exposure to inquiry, educational level, teaching experience, subject knowledge, and verbal, quantitative, and analytical reasoning skills. An instrument was used to quantify the effectiveness with which instructors use inquiry as an instructional technique. As expected, performance on tests of scientific reasoning and analytical reasoning skills were predictors of effective inquiry use. To test the second hypothesis on the relationship between the effective use of inquiry and pedagogical outcome, data on students' scientific reasoning skills, understanding of the nature of science, subject knowledge, and overall satisfaction with the instructor were gathered. As expected, students of instructors who used inquiry more effectively experienced greater normalized gains in scientific reasoning than students of instructors who used inquiry less effectively. A negative correlation between instructor inquiry use and student understanding of the nature of science was identified. No other student variables were significantly influenced by effectiveness of inquiry use. Seven appendixes contain student surveys and sample assessment items. (Contains 7 figures, 2 tables, and 51 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 456 157

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to
improve reproduction quality.
Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

Russell Benford

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Relationships Between Effective Inquiry Use and the Development of Scientific Reasoning Skills in College Biology Labs

Russell Benford
Anton E. Lawson

Department of Biology
Arizona State University
PO Box 87501
Tempe, AZ 85287

April 2001

TM033220

This material is based primarily on research partially supported by the National Science Foundation under Grant DUE 9453610. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

BEST COPY AVAILABLE

ABSTRACT

Two hypotheses regarding the relationship between scientific reasoning skills and the use of the inquiry method of instruction in college biology labs were examined. The first hypothesis was that scientific reasoning skills influence an instructor's ability to teach biology using inquiry. The second hypothesis was that the effectiveness with which an instructor uses inquiry affects the pedagogical outcome of a lesson. To test the first hypothesis, nine instructors teaching 702 students in an introductory biology course for non-majors were evaluated for their scientific reasoning skills and understanding of the nature of science. Data were also collected on instructors' prior exposure to inquiry, educational level, teaching experience, subject knowledge, and verbal, quantitative, and analytical reasoning skills. An instrument was used to quantify the effectiveness with which instructors use inquiry as an instructional technique. As expected, performance on tests of scientific reasoning and analytical reasoning skills were predictors of effective inquiry use. To test the second hypothesis on the relationship between the effective use of inquiry and pedagogical outcome, data on students' scientific reasoning skills, understanding of the nature of science, subject knowledge, and overall satisfaction with the instructor were gathered. As expected, students of instructors who used inquiry more effectively experienced greater normalized gains in scientific reasoning than students of instructors who used inquiry less effectively. A negative correlation between instructor inquiry use and student

understanding of the nature of science was identified. No other student variables were significantly influenced by effectiveness of inquiry use.

THEORETICAL RATIONALE

One goal of this study was to test the hypothesis that instructors' scientific reasoning skills influence their ability to effectively use inquiry as an instructional technique. In theory, an instructor's ability to lead an inquiry investigation in science should depend on familiarity with both the process and products of science, as well an understanding of the developmental psychology of students (see Figure 1). Presumably, effective inquiry instructors must demonstrate competency in their area of specialization as evidenced by their ability to provide accurate and meaningful descriptions and explanations, to help students identify

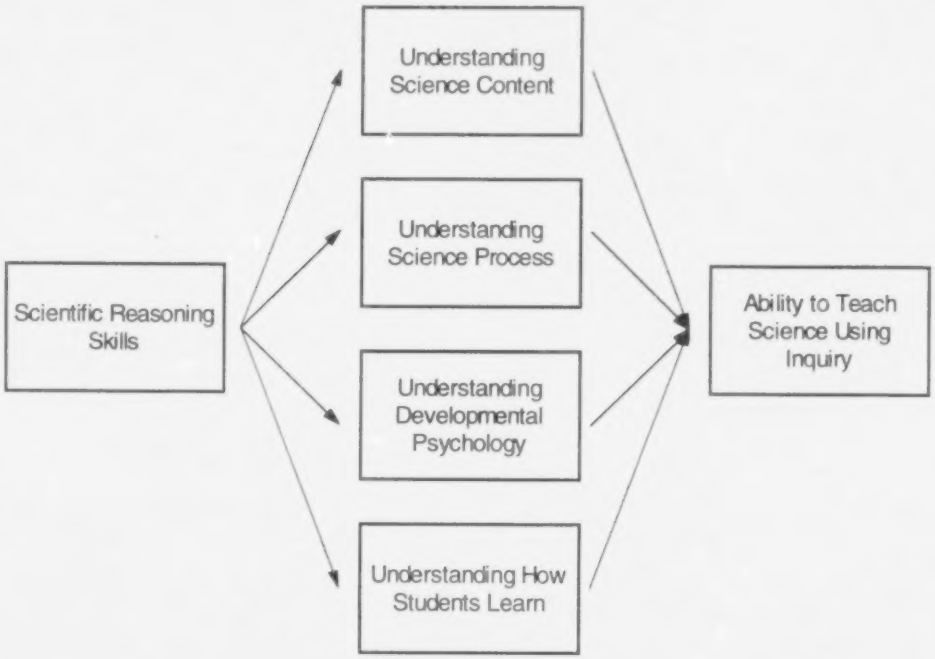


Figure 1. Theoretical relationships among scientific reasoning skills and the ability to teach science using inquiry

and avoid misconceptions, and to encourage students to make connections with other areas of knowledge. In addition to subject proficiency, effective inquiry instructors presumably need to understand the process of science and the nature of scientific inquiry well enough to facilitate student-centered investigations that involve exploring natural phenomena, identifying patterns, asking questions, generating and testing hypotheses, analyzing results, and accepting or rejecting proposed explanations based on an objective evaluation of empirical evidence.

But it is likely that instructors need more than a strong science background to succeed in an inquiry-oriented classroom. In theory, the skills to teach science using inquiry also depend on instructors' understanding of how and why inquiry is an effective pedagogical technique. Understanding how students construct knowledge based on personal experience, social interaction, and the analysis and interpretation of data is essential for effective inquiry instruction.

Inquiry science instructors must have the ability to challenge students cognitively and to be sensitive to their educational needs. Anticipating students' thoughts and behaviors, asking insightful and thought-provoking questions, engaging in pedagogically relevant discourse, and empathizing with students' frustrations and cognitive limitations are important skills associated with effective inquiry use.

Different instructors exhibit different patterns of cognition (Garnett & Tobin,1984; Lawrenz & Lawson, 1986; Lawson, 1999b). Individuals whose reasoning skills include the ability to categorize objects, events, and situations, to manipulate empirical variables, and to test categorical hypotheses can be classified as "concrete operational"

thinkers. Those whose reasoning skills include the ability to test causal hypotheses involving visible causal agents that correspond with the independent variable of the experiment can be classified as "formal operational." Those whose reasoning skills include the ability to test causal hypotheses involving unseen, theoretical causal agents that do not correspond with the independent variable of the experiment can be classified as "formal operational." (Lawson, Clark, Cramer-Meldrum, Falconer, Seaquist, & Kwon, 2000; Lawson, Drake, Johnson, Kwon, & Scarpone, 2000; Lawson, Alkhoury, Benford, Clark, & Falconer, 2000; and Lawson, 2001). Because, in theory, using inquiry effectively requires competency in all these areas of reasoning, formal and post-formal operational reasoning skills should be prerequisite for the effective use of the inquiry method of instruction in a science classroom.

To test this hypothesis, a measure of the scientific reasoning skills of nine instructors was administered before each instructor was trained and assigned to lead a semester of inquiry-oriented biology labs. Instructors were given a period of time to familiarize themselves with the inquiry method of instruction, then a measure of each instructor's ability to use inquiry was employed. If instructors' scientific reasoning skills do influence their ability to use inquiry effectively, then instructors who score higher on the test of scientific reasoning skills should demonstrate greater proficiency with inquiry than instructors who score lower on the test. Conversely, if effective inquiry use depends on other factors (e.g., prior exposure to inquiry, educational level, teaching experience, subject knowledge, or verbal, quantitative, or analytical reasoning skills), then instructors who perform better on measures of these factors should demonstrate

greater proficiency with inquiry than instructors who do not perform as well on measures of these variables.

In addition to testing the hypothesis that scientific reasoning skills influence the ability to use inquiry effectively, this study also tested the hypothesis that effective inquiry instruction contributes to the development of students' scientific reasoning skills. In theory, effective inquiry use in science exercises students' scientific reasoning skills by challenging students to explain natural phenomena by generating and testing hypotheses and by analyzing and interpreting data. Such an effort requires students to use correlational, combinatorial, proportional, and probabilistic reasoning, to identify and control variables, and, in some instances, to visualize unseen causal agents. Cognitive self-regulation, feedback, and constructive criticism from classmates and the instructor should help students confront errors and inconsistencies and rectify logical mistakes and scientific misconceptions. Thus, effective inquiry use should help students improve their scientific reasoning skills by increasing their familiarity with scientific problem-solving techniques and by encouraging deliberation and cognitive self-regulation.

To test this hypothesis, a measure of scientific reasoning skills was administered to students before and after the instructional treatment. If effective inquiry use improves students' scientific reasoning skills, then students of instructors who demonstrate greater proficiency with inquiry should have greater gains in scientific reasoning skills over the course of the semester than students of instructors who demonstrate less proficiency with the technique. Additionally, if effective inquiry use is better at producing gains in other domains (e.g., subject knowledge, understanding the nature of science,

satisfaction with the instructor), then students of instructors who use inquiry more effectively should perform better on tests of subject knowledge and understanding the nature of science, and/or give instructors higher satisfaction ratings than students of instructors who use the technique less effectively.

RELATED RESEARCH

What factors influence a science instructor's willingness and/or ability to use student-centered teaching practices such as inquiry? Numerous studies have attempted to answer this question by soliciting the opinions of instructors and school administrators using interviews, questionnaires, and classroom visits (e.g. Bainer, 1997; Costenson & Lawson, 1986; Lasley, Matczynski, & Benz, 1998; Loucks-Horsley, Stiles, & Hewson, 1996; Mittal, 1986; Sage & Torp, 1997; Staten, 1998; Sunal, 1975; Tamir, 1976; Tilgner, 1990; Tulloch, 1986; and Tuyay, Floriani, Yeager, Dixon, and Green, 1995). Results indicate that a large number of instructors and school administrators believe that affective and environmental factors such as administrative support, attitude, collaboration, confidence, experience, feedback, motivation, priority, reflectivity, and training influence an instructor's capacity and propensity to use student-centered teaching techniques. Costenson & Lawson (1986) also suggested that cognitive factors, including understanding the process of scientific inquiry and the structure of biology, might influence the effectiveness with which an instructor uses inquiry. Such

studies provide important insights into the minds of science instructors and school administrators, and they help us identify factors that might lead to effective inquiry use. However, additional research that more accurately defines the aforementioned terms and rigorously tests these hypotheses is necessary to identify the factors that contribute to inquiry use in the classroom.

Some workers have empirically tested hypotheses concerning the role of scientific reasoning skills in effective inquiry use. McKenna (1983) explored the relationship between the scientific reasoning skills of pre-service elementary school instructors and their propensity to use inquiry versus expository teaching methods. He found that pre-service elementary school instructors who are in the transitional stages of cognitive development display more inquiry-oriented behaviors than those in the concrete and formal operational stages. To explain the stronger inquiry orientation of transitional instructors, McKenna speculated that, while concrete operational instructors might "not have the cognitive abilities to use inquiry effectively," formal operational instructors might "lack understanding of the intellectual problems facing the children, probably because they can not relate to problems of abstract reasoning not encountered by themselves."

Lawrenz & Lawson (1986) investigated the relationship between the scientific reasoning skills of in-service elementary school instructors and student gains in scientific reasoning skills over one semester. They found that students of concrete operational elementary school instructors showed greater gains in scientific reasoning skills than students of their formal operational peers. Like McKenna (1983), Lawrenz &

Lawson hypothesized that students of concrete operational instructors showed greater gains in scientific reasoning because "instructors who were categorized as concrete operational ... were more sensitive to student difficulties than those who were categorized as formal operational because the concrete operational instructors think in ways more similar to their students." Lawrenz & Lawson also reported that students of instructors who stated a preference for using inquiry achieved slightly (but not significantly) greater gains in scientific reasoning skills than students of instructors who stated a preference for expository teaching methods.

In a subsequent study, Lawrenz (1988) investigated relationships between the scientific reasoning skills and teaching behaviors of elementary school instructors. While some of the results that Lawrenz obtained indicated that "concrete reasoners believed significantly more strongly in teaching specific science concepts than the... [teachers] classified as formal reasoners," other results indicated that "[there are] few consistent differences between the concrete and formal reasoners." In this study, Lawrenz found that the scientific reasoning skills of the instructors did not significantly influence their attitudes toward teaching or their teaching orientations. Thus, uncertainty exists about whether differences in attitudes and behaviors characterize instructors at various stages of cognitive development.

While McKenna's (1983), Lawrenz & Lawson's (1986), and Lawrenz' (1988) studies investigated relationships between instructor scientific reasoning skills and inquiry use, none documented a clear relationship between the two variables. Although these results could suggest that no such relationship exists, they could also suggest

that the instruments or methodologies used were not sensitive enough to detect a discrete relationship that *does* exist For example, McKenna (1983) used the Science Lesson Analysis System (Hacker, 1982) to rate the teaching orientations of his study participants, but he reported concerns about the reliability of the instrument. Additionally, McKenna's investigation focused on the teaching orientation of the study participants, but did not investigate pedagogical implications of the reported orientations. Lawrenz & Lawson (1986) relied on self-reporting to characterize participants' teaching orientations. And Lawrenz (1988) used several instruments to measure attitude and teaching orientation, which provided a variety of results that were sometimes contradictory. Thus, although previous research failed to document a clear link between instructor scientific reasoning skills, teaching orientation, and overall pedagogical effectiveness, there are reasons to further investigate the relationship between instructor scientific reasoning skills and effective inquiry use.

Numerous studies have demonstrated that inquiry-oriented teaching methods are more effective than traditional expository methods at improving students' attitudes toward science, content knowledge, and scientific reasoning skills (for reviews see Lawson, Abraham, & Renner, 1989; Lott, 1983; Shymansky, 1984). However, no previous studies have identified any specific factors that contribute to effective inquiry use. Additionally, no prior studies have documented factors that lead to or result from the diversity of teaching orientations or skills of science instructors. Prior research has investigated differences between expository and inquiry teaching approaches, but no

prior work has investigated the factors leading to or consequences of varying degrees of inquiry use by different instructors teaching the same lesson.

As a result, the present study will investigate factors that influence an instructor's ability to use inquiry effectively, as well as the pedagogical implications of effective inquiry use. In addition to scientific reasoning skills, factors that might influence an instructor's ability to use inquiry effectively include the level of higher education the instructor has attained, semesters of teaching experience, number of exposures to the inquiry approach in pre-service training, number of exposures to the inquiry approach in in-service training, understanding of the nature of science, subject knowledge, and verbal, quantitative, and analytical skills. Pedagogical implications of effective inquiry use will be investigated in the domains of student reasoning skills, subject knowledge, understanding of the nature of science, and overall satisfaction with the instructor.

Understanding factors that influence and even predict the effective use of inquiry in science classrooms could give science education researchers, administrators, and instructors the means to encourage and improve the use of inquiry in science education.

METHOD

Sample

Nine graduate teaching assistants (five males and four females, aged 22-34 years, mean age = 26.0 years) and 702 undergraduate students (442 females and 260 males, aged 17-58 years, mean age = 20.5 years) enrolled in a freshman level introductory biology course for non-majors at a large suburban university in the southwestern United States participated in the study. Each week students attended three 50-minute lectures delivered by the course professor. In addition to the lectures, students participated in a weekly two-hour lab. Each week each teaching assistant taught three lab sections enrolling approximately 25 students.

Design

Prior to the beginning of the semester, all teaching assistants participated in a three-day workshop that introduced the inquiry method of instruction (Lawson, et al., 1989). In the workshop, each teaching assistant was administered tests to measure their scientific reasoning skills and their understanding of the nature of science. Data regarding prior teaching experience, prior exposure to the inquiry method of instruction, educational background, verbal skills, quantitative skills, analytical skills, and general biological knowledge were also collected.

In the first week of the semester, students were also administered tests measuring their scientific reasoning skills and their understanding of the nature of science. Students then participated in a 15-week sequence of inquiry-oriented biology labs (Lawson, 1995) (See Table 1). The labs focused on conceptual understanding of natural phenomena and the development of scientific reasoning skills. Concepts the labs addressed included geologic time, natural selection, skull structure and function, behavioral ecology, photosynthesis, intraspecific variation, Mendelian genetics, function of invertebrate organ systems, biological communities, enzymatic reactions, osmosis, and air pressure. Cognitive skills the labs addressed included correlational, combinatorial, proportional, and probabilistic reasoning, causality, identification and control of variables, and the visualization of unseen causal agents.

Week	Name
1	What do fossils tell us about life in the past?
2	How do species evolve?
3	What can be inferred from animal structure?
4	Why don't birds get along?
5	What variables affect the rate of photosynthesis?
6	What causes intraspecific variation?
7	What determines specific characteristics in fruit flies?
8	What human characteristics covary?
9	What's inside a squid?
10	What variables affect heart rate?
11	How does the environment affect the distribution of organisms?
12	What happens during chemical reactions?
13	What variables affect the passage of molecules through cell membranes?
14	*No labs this week*
15	How can a burning candle cause water to rise?

Table 1. *Sequence of labs taught during study*

During the semester, teaching assistants participated in weekly two-hour meetings to discuss inquiry teaching methods and to prepare them for the next week's lab. Thus, teaching assistants were given repeated opportunities to improve their inquiry teaching skills over the course of the semester.

During that same period, students were given the opportunity to develop their reasoning skills, construct an understanding of the nature of science, and construct an understanding of various biological concepts by asking descriptive ("what") and causal ("why") questions about observed natural phenomena, by generating multiple hypotheses to attempt to answer the questions, by generating tests and predicting results, and by comparing predicted with actual results to support or reject their hypotheses.

The inquiry teaching skills of the teaching assistants were evaluated during the last lab of the semester, which was taught during the 15th week of instruction. This lab challenged students to investigate what happens when an inverted cylinder is placed over a burning candle sitting upright in a dish of water (Elementary Science Study 1974; Lawson, 1995; Lawson, 1999b; Lawson, Drake, Johnson, Kwon, & Scarpone, 2000). Students were encouraged to generate and test hypotheses to explain what causes water to rise in a cylinder when placed it is over the candle. Teaching assistants facilitated the investigation using the inquiry teaching skills they had developed over the course of the semester. Because having a meaningful conceptual understanding of air pressure requires the visualization of unseen, theoretical entities (rapidly moving air

molecules), ultimate success in this lab presumably required post-formal operational reasoning skills (Lawson, Alkhoury, Benford, Clark, & Falconer, 2000; Lawson, 2001).

The effectiveness with which each teaching assistant used inquiry was measured using the Reformed Teaching Observation Protocol (Sawada, Piburn, Falconer, Turley, Benford, & Bloom, 2000). Two independent observers evaluated each teaching assistant on separate days and in separate lab sections. The second and third of three lab sections in sequence were observed so that teaching assistants had an opportunity to facilitate the lesson once before their performance was evaluated. Teaching assistant RTOP scores were averaged, and the average scores were recorded.

At the end of the semester, students were administered a comprehensive final exam. The exam contained the questions from the Classroom Test of Scientific Reasoning and the Nature of Science Survey that had been administered at the beginning of the semester, so that gains in these areas of competency could be measured. The exam also contained questions to evaluate students' scientific reasoning skills and understanding of the nature of science in novel contexts. These questions were designed to measure the same cognitive skills that the Classroom Test of Scientific Reasoning and the Nature of Science Survey report to measure, but the questions were original so that prior exposure could not have influenced student responses. These questions were therefore considered scientific reasoning and nature of science transfer tests. Original questions testing students' comprehension of biological concepts introduced in labs and a survey to determine each student's overall satisfaction with their teaching assistant were also embedded in the final exam.

Normalized gains were calculated on the scientific reasoning and nature of science tests. Raw scores were used for the scientific reasoning transfer test, the nature of science transfer test, and subject knowledge test.

Instruments

Inquiry Teaching. To quantify the effectiveness with which a teaching assistant used inquiry in the classroom, the Reformed Teaching Observation Protocol (RTOP) (Sawada, et al., 2000) was used. The RTOP is a 25-item criterion-referenced observational instrument that quantifies the extent to which science and math teachers use inquiry techniques, as defined by the National Research Council (1990, 1995), the American Association for the Advancement of Science (1990, 1993), and the National Council for the Teaching of Mathematics (1989, 1991, 1995).

Using the RTOP, an observer assigns 0–4 points on each item relative to the absence or presence of 25 different instructor or student behaviors relating to questioning techniques, lesson design and implementation, locus of control, communicative interactions, and classroom culture. An overall score of 0–100 is awarded to the instructor, based on the sum of the points assigned for each item. Sample RTOP items include: "In this lesson, student exploration preceded formal presentation," "Students used a variety of means (models, drawings, graphs, concrete materials, manipulatives, etc.) to represent phenomena," "Students made predictions, estimations and/or hypotheses and devised means for testing them," "Student questions

and comments often determined the focus and direction of classroom discourse," and "The teacher acted as a resource person, working to support and enhance student investigations." A complete list of the items in the RTOP is shown in Appendix I. In previous studies, the RTOP has demonstrated a high inter-rater reliability (Cronbach's α = 0.95). Face and internal validity of the instrument were established by Sawada (1999).

Scientific Reasoning Skills. Scientific reasoning skills were measured using a modified version of the Classroom Test of Scientific Reasoning (Lawson, 1978). Validity of the original test was established by several studies (e.g. Lawson, 1978; 1979; 1980; 1982; 1983; 1992; Lawson & Weser, 1990; Lawson, Baker, DiDonado, Verdi, & Johnson, 1993). The modified test includes 24 multiple-choice questions that identify reasoning patterns associated with correlational reasoning, probabilistic reasoning, proportional reasoning, combinatorial reasoning, identification and control of variables, and hypothesis testing involving observable and unobservable entities. Validity of the modified version has been established by Lawson (1999b) and Lawson, Clark, Cramer-Meldrum, Falconer, Seaquist, & Kwon (2000). A complete list of the items in the Classroom Test of Scientific Reasoning is in Appendix II.

Nature of Science. Understanding of the nature of science was measured using the Nature of Science Survey (Lawson, 1999a). The survey contains 13 items that address understanding scientific methodology and epistemological issues such as the

value of scientific theories and the intellectual accessibility of facts and the truth. Responses are given on a five-point Likert scale ranging from "A" (strongly agree) to "E" (strongly disagree). Sample items include: "The central goal of science is to explain natural phenomena," "A hypothesis is an educated guess of what will be observed under certain conditions," "Hypotheses/theories cannot be proved to be true beyond any doubt," "A hypothesis that gains support becomes a theory," and "Scientific statements that are just a theory are of little value." A complete list of the items in the Nature of Science Survey is in Appendix III.

Other Teaching Assistant Variables. Other teaching assistant variables included overall score on the Graduate Record Examination (GRE) subject test for Biology (Educational Testing Service, 1997). The GRE is a 2 hour and 50 minute timed test containing approximately 200 multiple choice questions based on major areas of study in Biology such as genetics, cellular biology, molecular biology, organismal biology, ecology, and evolution. While actual test items were not available for review, sample items from the 1995 GRE subject test in Biology include: "All of the following are typical of the prophase stage of mitosis EXCEPT the (A) appearance of sister chromatids joined at the centromere (B) condensation of chromatin (C) disappearance of nucleoli (D) replication of DNA (E) migration of centrioles to the poles," "The concentration of which of the following in the blood primarily determines the metabolic rate in homeothermic (warm-blooded) animals? (A) Norepinephrine (B) Thyroxine (C) Corticosterone (D) Growth hormone (E) Glucagon," and "Which of the following traits

appeared earliest in the phylogenetic history of birds? (A) Jaws (B) Lungs (C) Stapes (D) Cochlea (E) Shelled egg" (Educational Testing Service, 1995). Many test items from the 1995 Biology GRE can be categorized in the knowledge and comprehension levels of Bloom's Taxonomy (Bloom, et al., 1956). Therefore, teaching assistants' GRE Biology score was used as a measure of their general biological content knowledge.

Teaching assistants' verbal, quantitative, and analytical skills were assessed by their performance on the general Graduate Record Examination (GRE) (Educational Testing Service, 1997). According to the Educational Testing Service, the General GRE is a computer-based measure of cognitive skills "that are acquired over a long period of time and that are not related to any specific field of study" (Educational Testing Service, 1999). The General GRE takes approximately 4 hours to complete. The test yields a verbal, a quantitative, and an analytical score.

The verbal GRE score purports to reflect one's ability to analyze and evaluate written material, analyze relationships among component parts of sentences, and recognize relationships among words and concepts. A sample question from the verbal section of the GRE is "Choose the word or set of words for each blank that *best* fits the meaning of the sentence: Although the Impressionist painters appeared to earlier art historians to be _____ in their methods, recent analyses of their brushwork suggest the contrary – that, in fact, their technique was quite _____. (A) unstudied ... sophisticated (B) idiosyncratic ... effective (C) eclectic ... naive (D) lax ... fashionable (E) careless ... unpremeditated" (Educational Testing Service, 2001).

The quantitative GRE score purports to reflect one's basic mathematical skills and the ability to reason quantitatively. A sample question from the quantitative section of the GRE is "The entries in a flower show competition are 2 orchids, 4 roses, 3 tulips, and 2 violets. If a first-prize selection consists of one flower from each of the four categories, how many different first-prize selections are possible? (A) 11 (B) 24 (C) 48 (D) 96 (E) 576" (Educational Testing Service, 2001).

The analytical GRE score purports to reflect one's ability to understand and deduce information from structured relationships, analyze and evaluate logical arguments, and identify hypotheses and plausible causal explanations (Educational Testing Service, 1999). A sample question from the analytical section of the GRE is "Seven meetings – J, K, L, M, N, O, and P – are to be scheduled, one on each day of a week that begins on Sunday. The following restrictions apply: Meeting J must take place on Sunday; meeting K must take place after both meeting L and meeting M; meetings N, O, and P must take place on three consecutive days, not necessarily in that order. If meeting O is on Saturday, then meeting K must take place on (A) Monday (B) Tuesday (C) Wednesday (D) Thursday (E) Friday" (Educational Testing Service, 2001).

To quantify experience, teaching assistants self-reported the number of exposures they had to the inquiry method of instruction. They reported both the number of times as an undergraduate student they were enrolled in a class taught with the inquiry method, and the number of times they were considered to be an instructor or teaching assistant in a class taught with the inquiry method. Teaching assistants also

reported the total number of semesters of teaching experience they had, regardless of the teaching methods they employed. In addition to experience, the level of education each teaching assistant had attained, based on degree held, was reported. Options included bachelor's degree, post-baccalaureate degree, and Master's degree (all teaching assistants in the study were enrolled in either Master's or doctoral programs).

Student Subject Knowledge. A 30-item true-false test was constructed by the researchers to assess student understanding of specific biological terms and concepts introduced during the semester. Questions were written at the knowledge and comprehension levels of Bloom's Taxonomy (Bloom et al., 1956). Sample items include: "Air pressure increases at higher altitude," "According to gene theory, gene pairs separate independently during zygote production," "Combustion produces water molecules," "Osmosis occurs only through living cell membranes," and "Photosynthesis generally is the reverse of cellular respiration." A complete list of the items on the test of student subject knowledge is in Appendix IV.

Scientific Reasoning in Novel Contexts. To determine whether the scientific reasoning abilities measured by the Classroom Test of Scientific Reasoning were transferable into novel contexts, eleven additional multiple choice questions to test students' scientific reasoning skills were administered to students at the end of the semester. Questions were written to test students' scientific process knowledge and

post-formal operational reasoning abilities (Lawson, Alkhoury, Benford, Clark, & Falconer, 2000; Lawson, Clark, Cramer-Meldrum, Falconer, Seaquist, & Kwon, 2000) by using hypothetico-deductive reasoning to reject hypotheses involving theoretical entities such as water molecules moving through cell membranes, air molecules and helium atoms pushing on the inside surface of a balloon, and "scent" molecules by which salmon navigate to their home streams to spawn. Questions involving the homing skills of salmon also tested students' abilities to use probabilistic reasoning and interpret data from a data table. A complete list of the items on the test of scientific reasoning in novel contexts, see Appendix V.

Nature of Science in Novel Contexts. To determine whether the scientific reasoning skills measured by the Nature of Science Survey were transferable into novel contexts, seven additional questions to test students' understanding of the nature of science were administered to students at the end of the semester. Questions were written to assess students' epistemology and understanding of the scientific process. As with the Nature of Science Survey, responses were given on a five-point Likert scale ranging from "A" (strongly agree) to "E" (strongly disagree). Statements to which students responded included: "Current scientific theories portray nature more accurately than those they replaced," "Scientists think that atoms exist because they have seen them through powerful microscopes," and "New discoveries depend mostly on luck." A complete list of the items on the test of nature of science in novel contexts, see Appendix VI.

Satisfaction with Instructor. Students' opinions of various teaching assistant behaviors were measured at the end of the semester using an anonymous survey. Twelve questions regarding teaching assistant knowledge and behavior were asked. Students responded on a five-point Likert scale, with "1" corresponding to the most positive response, and "5" corresponding to the most negative response. Items included: "Do you have confidence in the teaching assistant's knowledge of the subject?", "Does the teaching assistant encourage student response?", and "What overall grade would you give the teaching assistant?" A complete list of the items on the student survey is in Appendix VII.

RESULTS

Teaching Assistant RTOP Scores

The RTOP scores Evaluator A awarded the nine teaching assistants ranged from 42 – 88. The RTOP scores Evaluator B awarded the teaching assistants ranged from 31 – 93. The range of the averaged RTOP scores was from 42 – 90. Figure 2 shows the relationship between the RTOP scores the two evaluators awarded ($r = 0.90$, $p = 0.001$). This high degree of inter-reliability compares favorably with coefficients reported in previous studies (Sawada 1999).

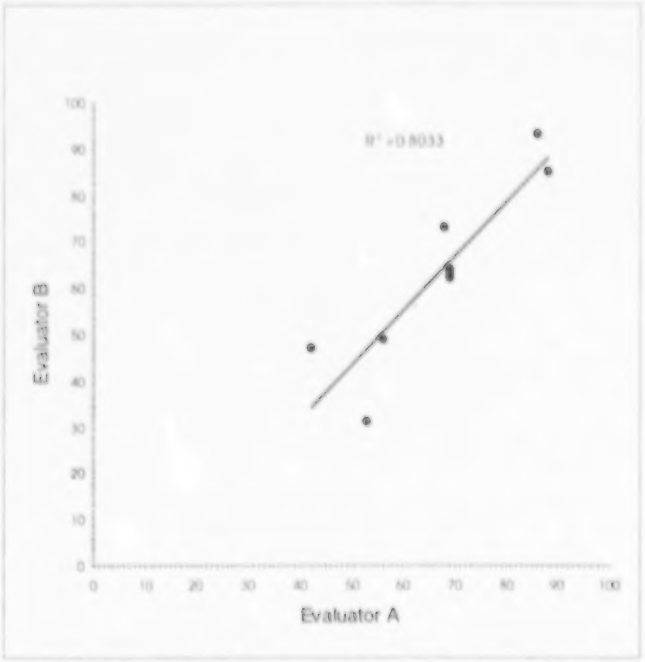


Figure 2. Inter-rater reliability of RTOP ($n = 9$, $r = 0.90$, $p = 0.001$)

To facilitate analysis in the second part of the study, teaching assistants were grouped into one of three categories based on their mean RTOP score: Low (mean RTOP scores of 42, 45, and 53), Medium (mean RTOP scores of 66, 66, 67, and 71), and High (mean RTOP scores of 87 and 90).

Predictors of Effective Inquiry Use During Instruction

Most teaching assistants had no prior exposure to inquiry teaching techniques. Because of the low variability in this category, this variable was eliminated from the

statistical analysis, and its potential influence on the teaching assistants' RTOP scores was not tested.

Instructors' level of education, semesters of teaching experience, understanding of the nature of science, verbal skills, quantitative skills, and subject knowledge, were not significant predictors of their RTOP scores. Interestingly, several variables correlated slightly [but not significantly] negatively with RTOP score: semesters of teaching experience ($n = 9$, $r = -0.17$, $p = 0.66$), verbal skills ($n = 9$, $r = -0.23$, $p = 0.56$), and subject knowledge ($n = 9$, $r = -0.22$, $p = 0.56$) (see Table 2).

The two remaining variables, performance on the test of scientific reasoning and performance on the analytical section of the GRE, correlated with teaching assistants' RTOP scores at levels approaching significance. The relationship between scores on the scientific reasoning test and RTOP scores ($r = 0.56$, $p = 0.12$) is illustrated in Figure 3. One teaching assistant was administered a slightly different version of the Classroom Test of Scientific Reasoning, so that person's score was eliminated from this analysis. The relationship between scores on the analytical section of the GRE test and RTOP scores ($r = 0.57$, $p = 0.14$) is illustrated in Figure 4.

To determine if the outlying point was responsible for the results of each analysis, data were re-analyzed excluding that point. Excluding the outlying data, the relationship between scores on the scientific reasoning test and RTOP scores ($r = 0.58$, $p = 0.18$), and the relationship between scores on the analytical section of the GRE and RTOP scores ($r = 0.60$, $p = 0.22$), were similar to the analyses that included the data, but they were not as strong.

Table 2
Correlation coefficients among study variables

	RTOP	Degree Earned	Teach. Exp.	Sci. Reas.	NOS Gain	GRE Verbal	GRE Quant.	GRE Analyt.	GRE Bio.
RTOP	1.00								
Degree Earned	0.34	1.00							
Teach. Exp.	-0.17	0.37	1.00						
Sci. Reas.	0.57	0.41	0.07	1.00					
NOS Gain	0.12	-0.48	-0.75**	-0.10	1.00				
GRE Verbal	-0.23	-0.04	0.01	-0.20	0.15	1.00			
GRE Quant.	0.33	-0.36	-0.71**	0.40	0.58	-0.25	1.00		
GRE Analyt.	0.56	-0.19	-0.82***	0.11	0.67**	-0.09	0.78	1.00	
GRE Bio.	-0.22	-0.26	0.39	0.33	-0.13	0.46	0.07	-0.29	1.00

Note. *p<0.10, **p<0.05, ***p<0.01. RTOP = RTOP Score; Teach. Exp. = Semesters of Teaching Experience; Sci. Reas. = Scientific Reasoning Skill; NOS Gain = Normalized Gain on Nature of Science Survey; GRE Quant. = GRE Quantitative; GRE Analyt. = GRE Analytical; GRE Bio. = GRE Biology

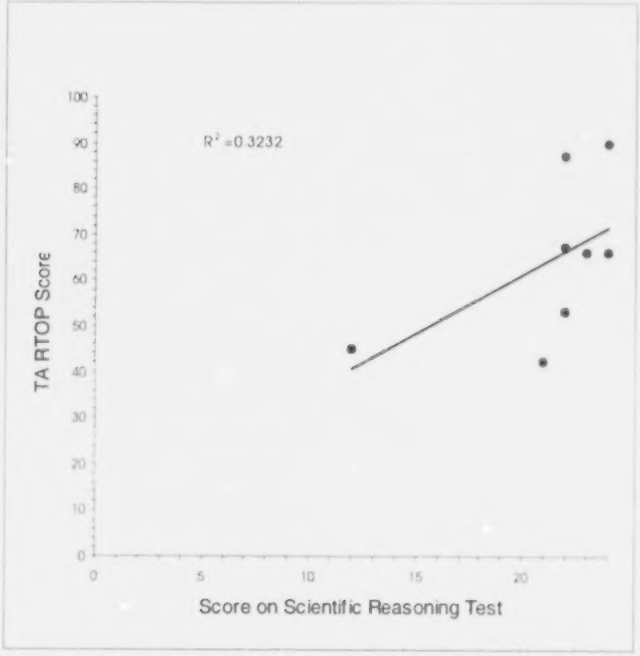


Figure 3. Correlation between teaching assistant RTOP score and score on test of scientific reasoning ($n = 8$, $r = 0.57$, $p = 0.14$)

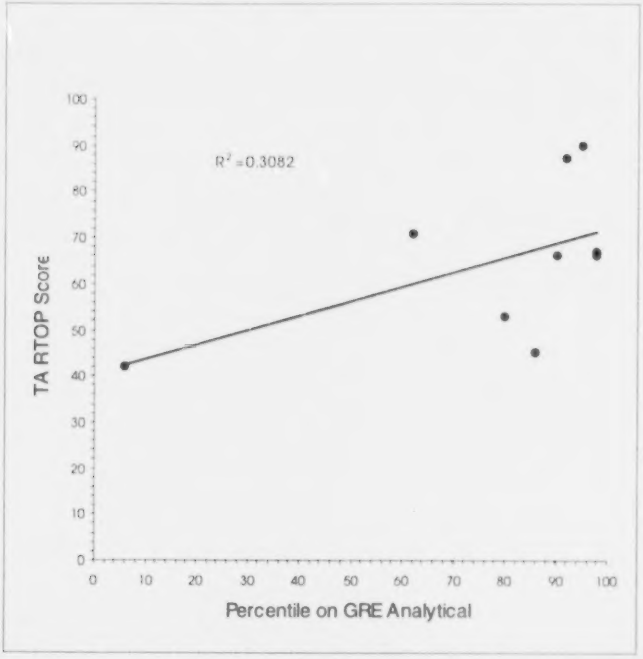


Figure 4. Correlation between teaching assistant RTOP score and score on analytical section of the GRE test ($n = 9$, $r = 0.56$, $p = 0.12$)

Consequences of Effective Inquiry Instruction

To investigate the pedagogical implications of effective inquiry instruction (i.e., high RTOP scores), students' pre-posttest gains in scientific reasoning and understanding the nature of science, as well as posttest scores on questions concerning the use of scientific reasoning abilities in novel contexts ("scientific reasoning transfer items"), applying an understanding of the nature of science in novel contexts ("nature of science transfer items"), and subject knowledge were analyzed.

While students of teaching assistants in all RTOP categories improved in scientific reasoning skills, a one-way analysis of variance showed that students of teaching assistants in different RTOP categories had significantly different normalized gains on the scientific reasoning test ($F_{2, 607} = 2.997$, $p = 0.05$). A post hoc Tukey's test characterized the difference between low and medium RTOP categories ($p = 0.07$), the low and high RTOP categories ($p = 0.10$), and the medium and high RTOP categories ($p = 0.99$). Figure 5 illustrates the relationship between the teaching assistant's RTOP category and normalized student gains on the scientific reasoning test.

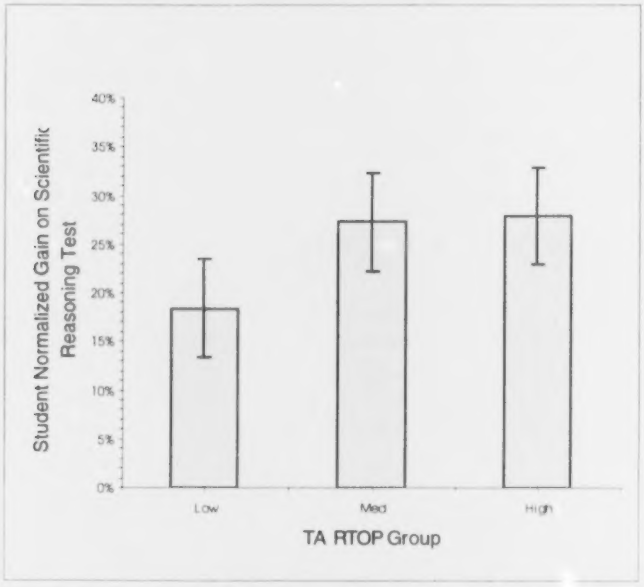


Figure 5. Normalized student gain on test of scientific reasoning by teaching assistant RTOP group ($F_{2, 607} = 2.997$, $p = 0.05$). Error bars represent +/- one standard deviation.

Students of teaching assistants in the low RTOP category showed a negligible gain in nature of science understanding; students of teaching assistants in the medium and high RTOP categories showed losses in that domain. A one-way analysis of

variance showed that students of teaching assistants in the three RTOP categories had significantly different normalized gains on the nature of science test ($F_{2, 573} = 3.416$, $p = 0.03$). A post hoc Tukey's test only revealed statistical significance between the low and medium RTOP categories ($p = 0.03$). Pairwise comparisons between the low and high RTOP categories (p = 0.14) and the medium and high RTOP categories (p = 0.94) were not significant. Figure 6 illustrates the relationship between the teaching assistant's RTOP category and normalized student gains on the nature of science test.

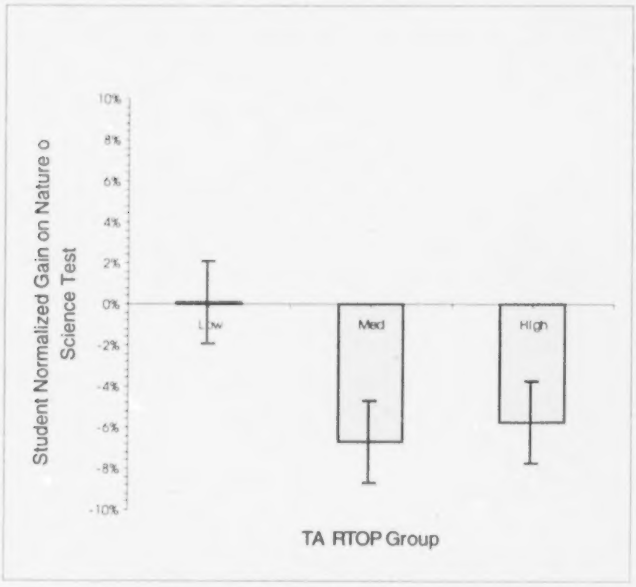


Figure 6. Normalized student gain on nature of science survey by teaching assistant RTOP group ($F_{2, 573} = 3.416$, $p = 0.03$). Error bars represent +/- one standard deviation.

No significant differences existed among students of teaching assistants in different RTOP categories on the scientific reasoning transfer test, the nature of science transfer test, and the subject knowledge test. No significant correlation existed between

teaching assistant RTOP score and student evaluation of teaching assistant (see Figure 7).

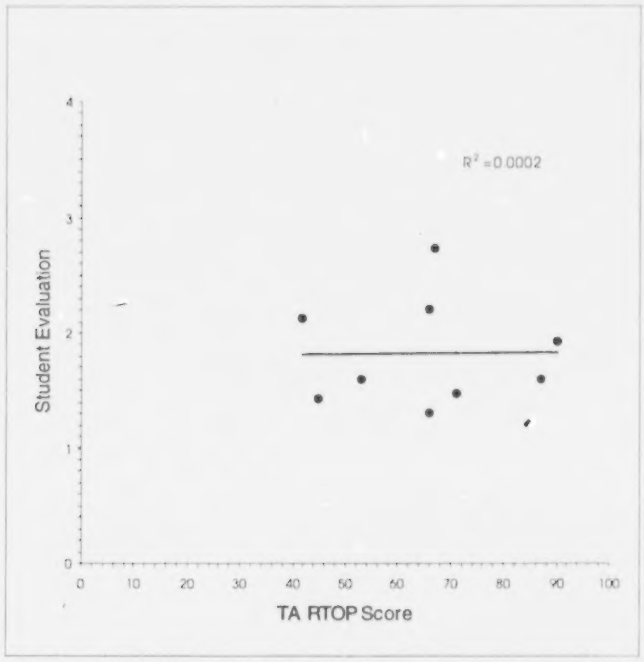


Figure 7. Student evaluation rating by teaching assistant RTOP score ($n = 9$, $r = 0.01$, $p = 0.97$)

DISCUSSION

As reported in Table 2 and Figure 3, I found a moderate correlation of 0.57 ($p = 0.14$) between teaching assistants' scientific reasoning skills and ability to use inquiry effectively in a college biology classroom. I predicted a positive correlation because, in theory, using inquiry effectively requires that teaching assistants understand complex theories of science and developmental psychology, and a meaningful understanding of

these theories presumably requires advanced scientific reasoning skills. Although the obtained correlation of 0.57 was not statistically significant, the sample size was small ($n = 8$). Figure 4 shows a similar correlation ($r = 0.56$) between teaching assistants' GRE Analytical Test scores and their RTOP scores. This was not surprising, since the Test of Scientific Reasoning and the GRE Analytical Test presumably measure related cognitive skills. Thus, these data provide some support for the hypothesis that teaching assistants' scientific reasoning skills influence their ability to lead an effective inquiry investigation in science. Additional research with larger sample sizes is necessary to further test this hypothesis.

The lack of support for alternative hypotheses that could explain the variation in effectiveness with which inquiry is used in college biology classrooms also contributes to the discussion on what factors influence or limit inquiry teaching success in general. As mentioned, several authors have reported that school administrators and instructors believe that scientific knowledge, laboratory skills, and instructional materials are limiting instructors' pedagogical success in student-centered classrooms. While the findings of this study do not directly refute these ideas, they do suggest that scientific reasoning skills could be another factor related to teacher success in inquiry-oriented science classrooms. Hone (1970) suggested that commonly cited affective and environmental factors are "science scarecrows" that are imaginary obstacles to the success of science teachers. These data provide some indirect support for Hone's hypothesis.

Figure 5 shows that teaching assistants who use inquiry more effectively produce greater gains in scientific reasoning skills in students than teaching assistants who use inquiry less effectively in a college-level biology labs. This positive correlation was also predicted, based on theory and the fact that previous research has shown that inquiry-oriented methodologies are more effective at producing gains in scientific reasoning abilities in students than traditional expository methodologies (cf. Lawson et al., 1989; Lott, 1983; Shymansky, 1984). The theoretical basis of inquiry instruction and the findings of these prior studies lead to the prediction that the stronger a teaching assistant's inquiry orientation, the greater gains in scientific reasoning skills that teaching assistant's students will realize. Data in this study are consistent with this prediction, and they show that teaching assistants who use inquiry more effectively (teaching assistants in the medium and high RTOP categories) produce greater gains in student scientific reasoning skills than teaching assistants who use inquiry less effectively (teaching assistants in the low RTOP category). Thus, these data provide additional support for the hypothesis that more effective inquiry use leads to greater gains in the scientific reasoning skills of college biology students.

Other authors who have investigated the pedagogical consequences of inquiry have reported at least superficially different results when working with differently aged students. As previously mentioned, Lawrenz & Lawson (1986) and McKenna (1983) found that concrete operational and transitional instructors using inquiry generate higher student achievement in elementary science classrooms than their formal operational colleagues. Attempting to explain this result, both parties hypothesized that the

similarities in thinking patterns among instructors and their [presumably] concrete operational and transitional students might be responsible for their success. If this hypothesis were correct, then it would also predict that teaching assistants who reason at a formal and post-formal operational levels should secure the most positive gains in students who reason at similarly advanced levels. In the present study, the teaching assistants who had more advanced scientific reasoning skills produced greater gains in the scientific reasoning skills of their students than the teaching assistants who had less advanced scientific reasoning skills. Thus, results from this study provide indirect support for Lawrenz', Lawson's, and McKenna's hypothesis that similarities between instructor and student thinking patterns might facilitate student success.

The combined results of Lawrenz & Lawson's (1986), McKenna's (1983), and this research leads to a broader hypothesis on the successful use of inquiry in science classrooms: inquiry-oriented science teachers who possess reasoning skills slightly more advanced than their students' reasoning skills are most effective pedagogically. This hypothesis seems reasonable, since facilitating student academic success might require strong enough intellectual skills to understand and teach the material, but also a degree of empathy for students' perspectives and means of understanding natural phenomena and scientific methodologies. In other words, the most effective science teachers might be those teachers who can stay a developmental step ahead of their students, but *not* those teachers who far outpace their students in cognitive ability. This hypothesis warrants further investigation.

Figure 6 shows that, on average, students of teaching assistants who use inquiry more effectively actually decrease more in their understanding of the nature of science than students of teaching assistants who use inquiry less effectively. This result is unexpected, and it weakens support for the general hypothesis that advanced scientific reasoning skills of instructors contribute to student performance in science.

Why does this negative correlation exist? Several hypotheses are forthcoming. First, effective inquiry use might not contribute to students' understanding of the nature of science, because effective inquiry use might not directly address this issue. Instead, more expository methods of teaching might be more successful at facilitating learning in this domain. Alternatively, students of effective inquiry instructors might experience an increased dissatisfaction with their own understanding of the nature of science and not be able to accommodate or resolve such cognitive disequilibrium in one semester. Or perhaps teaching assistants who reason with concrete operational and transitional patterns might strongly emphasize the terms and definitions associated with various steps of the scientific enterprise, whereas teaching assistants who reason at advanced levels of cognitive development might simply encourage students to explore and explain natural phenomena ... in other words, to *do* science instead of memorize terms associated with scientific philosophies and techniques. Anecdotal evidence from the classroom observations performed in this study lends support to this hypothesis, but further investigation is necessary to empirically test all the above hypotheses and explain this puzzling negative correlation.

CONCLUSION AND EDUCATIONAL IMPLICATIONS

In conclusion, the most educationally significant findings of this study are that the ability to use inquiry effectively in a college-level science labs seems to be most strongly associated with the formal and post-formal scientific reasoning skills of teaching assistants, and that the effective use of inquiry promotes the acquisition of formal and post-formal scientific reasoning skills in college-level students. No other variables, including experience and subject knowledge, predicted effective inquiry use among the teaching assistants. Thus, it seems that teaching assistants who possess advanced scientific reasoning skills have the greatest chance of helping their students acquire scientific reasoning skills using inquiry as a pedagogical tool.

With this in mind, Garnett and Tobin's (1984) finding that "large numbers of preservice instructors do not use formal reasoning patterns" is of particular concern. If science instructors have underdeveloped scientific reasoning skills, then their ability to guide inquiry lessons involving complex scientific theories and improve the advanced scientific reasoning skills of their students might be suspect. It seems logical that elementary teachers who facilitate the development of concrete operational student skills such as categorization and the testing of categorical hypotheses might produce appreciable student outcomes if they reason at concrete operational or transitional stages. However, secondary and post-secondary teachers who facilitate the development of formal and post-formal operational student cognitive skills involving

unseen, theoretical entities and independent variables that do not directly correspond with hypothesized causal agents might only produce appreciable student outcomes if they reason at the post-formal level of cognitive development.

The challenge of educating science teachers is multi-faceted, and the breadth of knowledge and experience a new teacher must acquire to ensure her or his effectiveness is considerable. But, if, as numerous national science advocacy organizations recommend, the acquisition of advanced scientific reasoning skills is a primary goal of science education, and, as this and other research suggests, effective inquiry is a useful vehicle for developing those skills, then teacher education curricula that focus primarily on the development of teachers' scientific reasoning skills should have strong and measurable impacts on student achievement in secondary and post-secondary science.

REFERENCES

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

American Association for the Advancement of Science. (1990). *Science for All Americans*. New York: Oxford University Press.

Bainer, D.L. (1997). With a new lens: How partnering impacts teaching assistants' views of and approaches to teaching science. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Bloom, B.S. et al. (1956). Taxonomy of educational objectives; the classification of educational goals, by a committee of college and university examiners. New York: Longmans, Green.

Costenson, K. & Lawson, A.E. (1986). Why isn't inquiry used in more classrooms? *The American Biology Teacher*, 48, 150-158.

Educational Testing Service. (1995). *Practicing to take the biology test: 3rd Edition*. Princeton, NJ: Author.

Educational Testing Service. (1997). *Graduate record examination*. Princeton, NJ: Author.

Educational Testing Service. (1999). "Graduate record examinations information and registration bulletin." Princeton, NJ: Author.

Educational Testing Service. Mar 6, 2001 <http://grediagnostic.ets.org>.

Elementary Science Study. (1974). *Gasses and Airs: Instructor's Guide*. New York: McGraw Hill.

Garnett, P.J., & Tobin, K. (1984). Reasoning patterns of preservice elementary and middle school science teaching assistants. *Science Education*, 68, 621-631.

Hacker, R.C. (1982). The science lesson analysis system (Research Report No. 1). Perth: The Centre for Science Education. The University of Western Austrailia.

Hone, E. (1970). Science scarecrows. *School Science and Mathematics*, 70, 322-326.

Lasley, T.J., Matczynski, T.J., & Benz, C.R. (1998). Science teaching assistants as change-ready and change-resistant agents. *The Educational Forum*, 62, 120-130.

Lawrenz, F. (1988). The relationships among elementary school teachers' attitudes and beliefs about science and their teaching style. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (61st, Lake of the Ozarks, MO, April 10-13, 1988).

Lawrenz, F. & Lawson, A.E. (1986). Student gain in reasoning ability as a function of teacher reasoning level and teaching style preference. *Journal of Research in Science Teaching*, 23, 523-531.

Lawson, A.E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15, 11-24.

Lawson, A.E. (1979). Relationships among performances on group-administered items of formal reasoning. *Perceptual and Motor Abilities*, 48, 71-78.

Lawson, A.E. (1980). The relationship among levels of intellectual development, cognitive style, and grades in a college biology course. *Science Education*, 64, 95-102.

Lawson, A.E. (1982). The reality of general cognitive operations. *Science Education*, 66, 29-241.

Lawson, A.E. (1983). Predicting science achievement: The role of developmental level, mental capacity, prior knowledge and beliefs. *Journal of Research in Science Teaching*, 20, 117-129.

Lawson, A.E. (1992). What do tests of formal reasoning actually measure? *Journal of Research in Science Teaching*, 29, 965-984.

Lawson, A.E. (1995). *Exploring the Living World: A Laboratory Manual for Biology (Instructor's Materials)*. New York: McGraw-Hill, Inc.

Lawson, A.E. (1999a). What should students know about the nature of science and how should we teach it? *Journal of College Science Teaching*, 28, 401-411.

Lawson, A.E. (1999b). Developmental level and overcoming nature of science misconceptions among pre-service biology teaching assistants. *Journal of Science Teacher Education* 23, 543-556.

Lawson, A.E. (2001). A theory of scientific argumentation and its implications for science teaching. *International Journal of Science Education*, submitted.

Lawson, A.E. & Weser, J. (1990). The rejection of nonscientific beliefs about life: The effects of instruction and reasoning abilities. *Journal of Research in Science Teaching*, 27, 589-606.

Lawson, A.E., Abraham, M.R., & Renner, J.W. (1989). *A theory of instruction: Using the learning cycle to teach science concepts and thinking abilities.* Cincinnati, OH: National Association for Research in Science Teaching.

Lawson, A.E., Alkhoury, S., Benford, R., Clark, B.R., & Falconer, K.A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching*, 37, 996-1018.

Lawson, A.E., Baker, W.P., DiDonato, L., Verdi, M.P. & Johnson, M.A. (1993). The role of physical analogies of molecular interactions and hypothetico-deductive reasoning in conceptual change. *Journal of Research in Science Teaching*, 30, 1073-1086.

Lawson, A.E., Clark, B., Cramer-Meldrum, E., Flaconer, K.A., Seaquist, J.M., & Kwon, Y.J. (2000). The development of scientific reasoning abilities in college biology: Do two levels of hypothesis-testing abilities exist? *Journal of Research in Science Teaching*, 37, 81-101.

Lawson, A.E., Drake, N., Johnson, J., Kwon, Y.J., & Scarpone, C. (2000). How good are students at testing alternative explanations of unseen entities? *American Biology Teacher*, 62, 249-255.

Lott, G.W. (1983). The effect of inquiry teaching and advance organizers upon student outcomes in science education. *Journal of Research in Science Teaching*, 20, 437-451.

Loucks-Horsley, S., Stiles, K., & Hewson, P. (1996). Principles of effective professional development for mathematics and science education: A synthesis of standards. NISE Brief, 1(1), 1-6.

McKenna, L.N. (1983). Relationships between cognition of teaching assistants and quality of teaching style in elementary science. Paper presented at the Congress of the Australian and New Zealand Association for the Achievement of Science, Perth, Australia.

Mittal, C.L. (1986). A comparison of educators' preconceptions of the competencies of secondary science teaching assistants in India. *Dissertation Abstracts International* 46(12A): 3691.

National Council for the Teaching of Mathematics (1989). *Curriculum and evaluation standards for school mathematics.* Reston, Virginia: The Council.

National Council for the Teaching of Mathematics (1991). *Professional standards for teaching mathematics.* Reston, Virginia: The Council.

National Council for the Teaching of Mathematics (1995). *Assessment standards for school mathematics.* Reston, Virginia: The Council.

National Research Council (1990). *National Science Education Standards.* Washington, DC: National Academy Press.

National Research Council (1990). *Fulfilling the Promise: Biology Education in the Nation's Schools.* Washington, DC: National Academy Press.

Sage, S.M., & Torp, L.T. (1997). What does it take to become a teaching assistant of problem-based learning? *Journal of Staff Development*, 18, 32-36.

Sawada, D. (1999). *Psychometric Properties of RTOP.* Technical Report No. IN99-2. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Instructors.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed Teaching Observation Protocol (RTOP)* (ACEPT Technical Report No. IN00-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Instructors.

Shymansky, J. (1984). BSCS programs: Just how effective were they? *The American Biology Teacher*, 46, 54-57.

Staten, M.E. (1998). Action research study: A framework to help move teaching assistants toward an inquiry-based science teaching approach. Milwaukee, Wisconsin: Milwaukee Public Schools.

Sunal, D. (1975). Affective measures for science teaching: Measures of teaching assistant interest and attitude toward science teaching. Morgantown, West Virginia: West Virginia University.

Tamir, P. (1976). "Investigations to inquiry" and teaching assistant training. The American Biology Teacher, 38, 50-52.

Tilgner, P.J. (1990). Avoiding science in the elementary school. *Science Education* 74, 421-431.

Tulloch, B.R. (1986). A factor analytic study of secondary science teaching assistant competencies within which growth is perceived as important by science teaching assistants, supervisors, and teaching assistant educators. *Journal of Research in Science Teaching*, 23, 543-556.

Tuyay, S., Floriani, A., Yeager, B., Dixon, C., & Green, J. (1995). Constructing an integrated, inquiry-oriented approach in classrooms: A cross case analysis of social, literate, and academic practices. *Journal of Classroom Interaction*, 30, 1-15.

APPENDIX I

REFORMED TEACHING OBSERVATION PROTOCOL (RTOP)

The following 25 items are behaviors associated with in inquiry science classroom, taken from the Reformed Teaching Observation Protocol (Sawada et al., 2000). This instrument quantifies the extent to which a classroom is an inquiry-oriented science classroom. To use the instrument, an independent observer awards 0 ("Never Occurred") – 4 ("Very Descriptive") points for each behavior. The total score (out of a total possible 100 points) gives the researcher an index with which she or he can measure an instructor's teaching orientation. Instructors with higher scores are considered to be more inquiry-oriented.

1. The instructional strategies and activities respected students' prior knowledge and the preconceptions inherent therein.

2. The lesson was designed to engage students as members of a learning community.

3. In this lesson, student exploration preceded formal presentation.

4. This lesson encouraged students to seek and value alternative modes of investigation or of problem solving.

5. The focus and direction of the lesson was often determined by ideas originating with students.

6. The lesson involved fundamental concepts of the subject.

7. The lesson promoted strongly coherent conceptual understanding.

8. The teacher had a solid grasp of the subject matter content inherent in the lesson.

9. Elements of abstraction (i.e., symbolic representations, theory building) were encouraged when it was important to do so.

10. Connections with other content disciplines and/or real world phenomena were explored and valued.

11. Students used a variety of means (models, drawings, graphs, concrete materials, manipulatives, etc.) to represent phenomena.

12. Students made predictions, estimations and/or hypotheses and devised means for testing them.

13. Students were actively engaged in thought-provoking activity that often involved the critical assessment of procedures.

14. Students were reflective about their learning.

15. Intellectual rigor, constructive criticism, and the challenging of ideas were valued.

16. Students were involved in the communication of their ideas to others using a variety of means and media.

17. The teacher's questions triggered divergent modes of thinking.

18. There was a high proportion of student talk and a significant amount of it occurred between and among students.

19. Student questions and comments often determined the focus and direction of classroom discourse.

20. There was a climate of respect for what others had to say.

21. Active participation of students was encouraged and valued.

22. Students were encouraged to generate conjectures, alternative solution strategies, and ways of interpreting evidence.

23. In general the teacher was patient with students.

24. The teacher acted as a resource person, working to support and enhance student investigations.

25. The metaphor "teacher as listener" was very characteristic of this classroom.

APPENDIX II

CLASSROOM TEST OF SCIENTIFIC REASONING

1. Suppose you are given two pieces of bread dough of equal size and shape. The two pieces also weigh the same. One piece is rolled up into a ball. The other is flattened into a pancake-shaped piece. *Which of these statements is correct?*

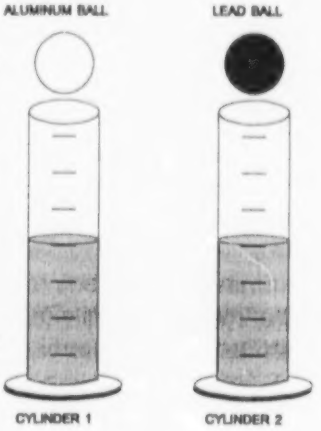
   a. The two pieces still weigh the same
   b. The ball of dough weighs more than the pancake-shaped piece
   c. The pancake-shaped piece weighs more than the ball

2. *because*

   a. when something is flattened it loses weight.
   b. dough has not been added or taken away.
   c. the flattened piece covers a larger area.
   d. the ball pushes down more on one spot.
   e. when something is flattened it gains weight.

3. To the right are drawings of two cylinders filled to the same level with water. The cylinders are identical in size and shape. Also shown are two metal balls, one made of aluminum and one made of lead. The metal balls are the same size but the lead ball is much heavier than the aluminum one.

   When the aluminum ball is lowered into Cylinder 1, it sinks to the bottom and the water level rises to the 4th mark. *If we lower the lead ball into Cylinder 2, the water will rise*

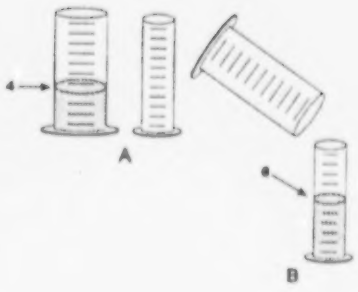


   a. to a higher level than it did in Cylinder 1
   b. to a lower level than it did in Cylinder 1
   c. to the same level as it did in Cylinder 1

4. *because*

   a. the lead ball is heavier than the aluminum ball.
   b. the balls are the same size.
   c. the lead ball will sink faster.
   d. the balls are made of different materials.

5. To the right are drawings of a wide and a narrow cylinder. The cylinders have equally spaced marks on them. Water is poured into the wide cylinder up to the 4th mark (see A). This water rises to the 6th mark when poured into the narrow cylinder (see B).



Both cylinders are emptied (not shown) and water is poured into the wide cylinder up to the 2nd mark. *How high would this water rise if it were poured into the empty narrow cylinder?*

a. to about 1
b. to about 2
c. to about 3
d. to about 4
e. none of these answers is correct

6. *because*

a. it went up 2 more before, so it will go up 2 more again.
b. the second cylinder is narrower.
c. one must actually pour the water and observe to find out.
d. the answer can not be determined with the information given.
e. 4 to 6 is the same ratio as 2 to 3.

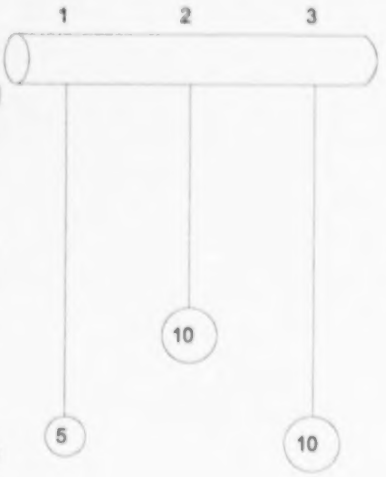
7. Water is now poured into the narrow cylinder (described in Item 5 above) up to the 7th mark. *How high would this water rise if it were poured into the empty wide cylinder?*

a. to about 6
b. to about 4
c. to about 5
d. to about 4½
e. none of these answers is correct

8. *because*

a. you subtract 2 from the wide for every 3 from the narrow.
b. it was 2 less before so it will be 2 less again.
c. the answer can not be determined with the information given.
d. the ratios must stay the same.
e. one must actually pour the water and observe to find out.

9. At the right are drawings of three strings hanging from a bar. The three strings have metal weights attached to their ends. String 1 and String 3 are the same length. String 2 is shorter. A 5 unit weight is attached to the end of String 1. A 10 unit weight is attached to the end of String 2. A 10 unit weight is also attached to the end of String 3. The strings (and attached weights) can be swung back and forth and the time it takes to make a swing can be timed.



Suppose you want to find out whether the length of the string has an effect on the time it takes to swing back and forth. *Which strings would you use to find out?*

a. 2 and 3
b. 1 and 3
c. 1 and 2
d. all three strings
e. only one string

10. *because*

a. the weights differ.
b. you must compare strings with both heavy and light weights.
c. to make all possible comparisons.
d. you must use the longest strings.
e. only the lengths differ.

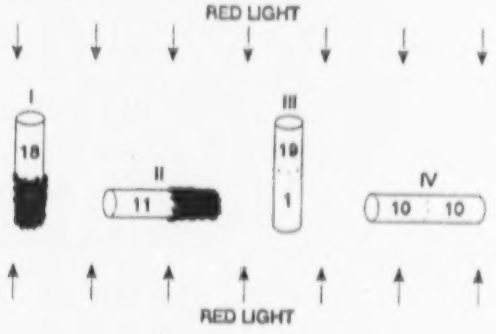
11. Suppose you want to find out whether the amount of weight attached to the end of a string has an effect on the time it takes for a string to swing back and forth. *Which of the strings in Item 9 above would you use to find out?*

    a. 2 and 3
    b. 1 and 3
    c. 1 and 2
    d. all three strings
    e. only one string

12. *because*

    a. to make all possible comparisons.
    b. you must use the heaviest weights.
    c. you must compare both long and short strings.
    d. only the weights differ.
    e. the lengths differ.

13. Twenty fruit flies are placed in each of four glass tubes. The tubes are sealed. Tubes I and II are partially covered with black paper; Tubes III and IV are not covered. The tubes are placed as shown. Then they are exposed to orange light for five minutes. The number of flies in the uncovered part of each tube is shown in the drawing.



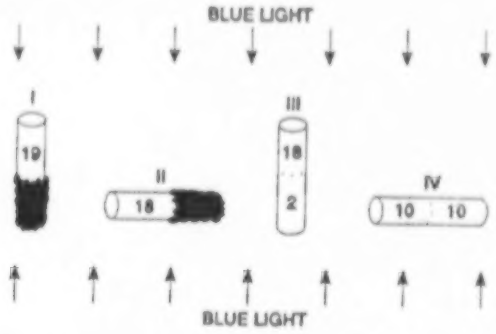
This experiment shows that flies respond to (respond means move to or away from):

    a. orange light but not gravity
    b. gravity but not orange light
    c. both orange light and gravity
    d. neither orange light nor gravity

14. *because*

a. some flies are in both ends of each tube.
b. the majority of flies are in the lighted ends and the lower ends of the tubes.
c. most flies went to the bottom of Tubes I and III.
d. the flies need light to see and must fly against gravity.
e. most flies are in the lighted end of Tube II but spread about evenly in Tube III.

15. In a second experiment, a different kind of fly and green light was used. The results are shown in the drawing.



These data show that these flies respond to (respond means move to or away from):

a. green light but not gravity
b. gravity but not green light
c. both green light and gravity
d. neither green light nor gravity

16. *because*

a. some flies are in both ends of each tube.
b. the flies are spread about evenly in Tube IV and in the lower end of Tube III.
c. most flies are in the lower end of Tube III and in the dark end of Tube II.
d. most flies are in the dark end of Tube II and in the upper end of Tube III.
e. the flies need light to see and must fly against gravity.

17. Eight triangular pieces of wood are put into a cloth bag and mixed about. The eight pieces are identical in size and shape, however, four pieces are black and four are white. Suppose someone reaches into the bag (without looking) and pulls out one piece. *What are the chances that the piece is black?*

    a. 1 chance out of 8
    b. 1 chance out of 1
    c. can not be determined
    d. 1 chance out of 2
    e. 1 chance out of 4

18. *because*

    a. only 1 black piece can be picked from the 4 black pieces.
    b. all 8 pieces are identical in size and shape.
    c. there is no way to tell which piece will be picked.
    d. only 1 piece of the 8 in the bag is picked.
    e. 4 out of 8 pieces are black.

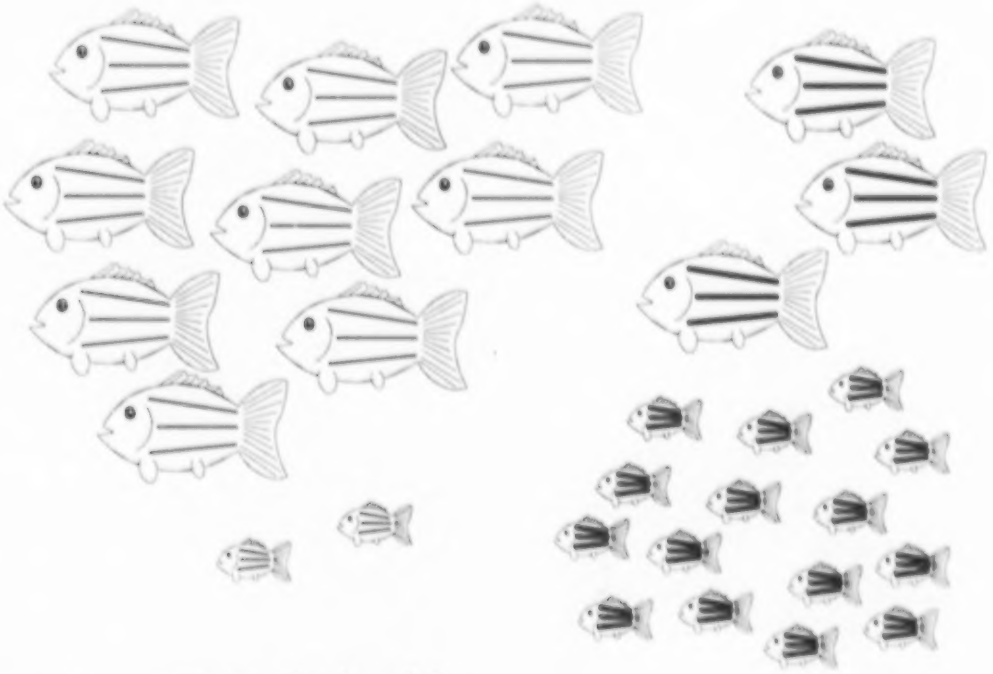
19. One red square pieces of wood, six yellow square pieces, and eight blue square pieces are put into a cloth bag. Two red round pieces, one yellow round pieces, and three blue round pieces are also put into the bag. All the pieces are then mixed about. Suppose someone reaches into the bag (without looking and without feeling for a particular shape piece) and pulls out one piece. *What are the chances that the piece is a yellow square or blue square piece?*

    a. 2 chances out of 3
    b. 1 chance out of 21
    c. can not be determined
    d. 1 chance out of 2
    e. 18 chances out of 21

20. *because*

a. only 1 of the 21 pieces is picked out of the bag.
b. 2 of every 3 pieces is a yellow or blue square piece.
c. 1 of the 2 shapes is square.
d. 18 of the 21 pieces are yellow or blue.
e. there is no way to tell which piece will be picked.

21. Look at the fish below that were caught by a fisherman one morning. The fisherman noticed that some of the fish were big and some were small. Also some had wide stripes and others had narrow stripes. This made the fisherman wonder if there might be a link between the size of the fish and the width of their stripes. *Do you think there is a link between the size of the fish and the width of their stripes?*

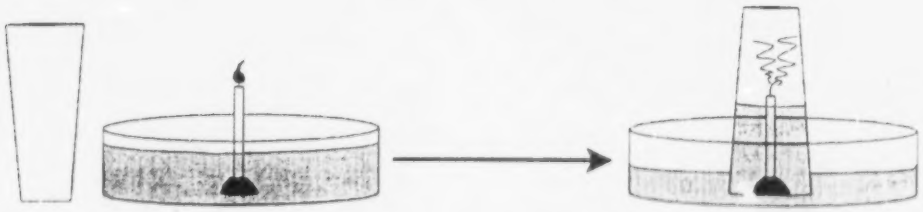


a. appears not to be a link
b. appears to be a link
c. can not make a reasonable guess

22. *because*

a. as the fish grow longer, their stripes become narrower.
b. there are some of each kind of fish.
c. there may be a genetic link between fish size and width of stripes.
d. most of the big fish have narrow stripes while most of the little fish have wide stripes.
e. there were not enough fish caught.

23. The figure below shows a drinking glass and burning birthday candle stuck in a small piece of clay standing in a pan of water. When the glass is turned upside down, put over the candle and placed in the water, the candle quickly goes out and water rushes up into the glass (as shown below at the right).



This observation raises an interesting question: Why does the water rush up into the glass?

Here is a possible explanation. The flame converts oxygen from the air into carbon dioxide. Because oxygen molecules do not dissolve very rapidly in water, but carbon dioxide molecules do, the newly-formed carbon dioxide molecules dissolve rapidly into the water lowering the air pressure inside the glass. Thus, the relatively higher air pressure outside the glass pushes the water up.

Suppose you have the materials mentioned above plus some matches and some dry ice (dry ice is frozen carbon dioxide). Using some or all of these materials, how could you best test this possible explanation?

a. Saturate the water with carbon dioxide and redo the experiment noting amount of water rise.
b. The water rises because oxygen is consumed; so redo the experiment in exactly the same way to show water rise due to oxygen loss.
c. Conduct a controlled experiment varying only the number of candles to see if that makes a difference.
d. Suction is responsible for water rise; so put a balloon over the top of an open- ended cylinder and place the cylinder over the burning candle.
e. Redo the experiment but make sure it is controlled by holding all independent variables constant; then carefully measure amount of water rise.

24. What result of your test (mentioned in item 23 above) would show that the explanation is probably wrong?

a. The water rises higher than it did before.
b. The water rises the same as it did before.
c. The water rises less than it did before.
d. The balloon expands out.
e. The balloon is sucked in.

APPENDIX III

NATURE OF SCIENCE SURVEY

Next to each item write the number that best reflects your current belief:
1=strongly disagree 2=disagree 3=don't know 4=agree 5=strongly agree

1. The central goal of science is to explain natural phenomena.
2. Hypotheses are derived from controlled observations of nature.
3. A hypothesis is an educated guess of what will be observed under certain conditions.
4. A conclusion is a statement of what was observed in statement number 3 above.
5. Hypotheses/theories cannot be proved to be true beyond any doubt
6. Hypotheses/theories can be disproved beyond any doubt.
7. To be scientific, a hypothesis must be testable.
8. To be tested, hypotheses must lead to expected results.
9. A hypothesis that gains support becomes a theory.
10. A theory that gains support becomes a law.
11. Truth is attainable via proof through repeated supporting observations.
12. The central goal of science is to discover facts about nature.
13. Scientific statements that are just a theory are of little value.

APPENDIX IV

STUDENT SUBJECT KNOWLEDGE

a = true b = false

1. Air pressure increases with increasing air temperature.

2. Air pressure decreases with decreasing volume.

3. Air pressure is caused by molecular collisions.

4. Air pressure is produced because nature "abhors" a vacuum.

5. Air pressure increases at higher altitude.

6. According to gene theory, sexually reproducing organisms have at least one pair of genes for each observable characteristic.

7. According to gene theory, one gene of a pair can dominate the expression of the other gene.

8. According to gene theory, gene pairs separate independently during zygote production.

9. According to gene theory, gene pairs recombine randomly during egg and sperm production.

10. According to gene theory, genes are located in chromosomes.

11. According to kinetic-molecular theory, a flame consists of rapidly moving particles that transfer energy to near by particles.

12. Molecules can combine or break apart when they collide.

13. Smaller molecules usually contain more chemical energy than larger molecules.

14. Fast moving molecules usually contain more kinetic energy than slow moving molecules.

15. Flame in an enclosed jar reduces the number of gas molecules in the jar.

16. Combustion (burning) "releases" stored chemical energy.

17. Combustion destroys oxygen atoms.

18. Combustion in an enclosed system increases the total amount of energy in that system.

19. Combustion produces water molecules.

20. Combustion utilizes chemical energy as its energy source.

21. Osmosis occurs when a dialysis bag filled with distilled water is placed in a glucose solution.

22. Osmosis is not effected by temperature.

23. Osmosis will not occur when membranes block diffusion.

24. Osmosis occurs only through living cell membranes.

25. Osmosis involves random ionic and/or molecular collisions.

26. Photosynthesis is carried out by green plants, mushrooms and cyanobacteria.

27. Photosynthesis actively transports glucose molecules into cell chloroplasts.

28. Photosynthesis involves light-capturing pigment molecules.

29. Photosynthesis uses solar energy to combine CO_2 with H_2O molecules to produce carbohydrate and O_2 molecules.

30. Photosynthesis generally is the reverse of cellular respiration.

APPENDIX V

SCIENTIFIC REASONING IN NOVEL CONTEXTS

Items 1 & 2 are based on the following information:

A student put a drop of blood on a microscope slide and then looked at the blood under a microscope. As you can see in the diagram below, the magnified red blood cells look like little round balls. After adding a few drops of salt water to drop of blood, the student noticed that the cells appeared to become smaller.

Magnified Red Blood Cells After Adding Salt Water

This observation raises an interesting question: Why do the red blood cells appear smaller?

Here are two possible explanations: *I. Salt ions (Na^+ and Cl^-) push on the cell membranes and make the cells appear smaller. II. Water molecules are attracted to the salt ions so the water molecules move out and leave the cells smaller.* To test these explanations the student used some salt water, a very accurate weighing device, and some water-filled plastic bags and assumed that the plastic behaves just like red-blood-cell membranes. The experiment involved carefully weighing a water-filled bag and placing the bag in a salt solution for ten minutes and reweighing the bag.

1. What result of the experiment would best show that explanation I is probably wrong?

   a. the bag loses weight
   b. the bag weighs the same
   c. the bag appears smaller

2. What result of the experiment would best show that explanation II is probably wrong?

   a. the bag loses weight
   b. the bag weighs the same
   c. the bag appears smaller

Items 3 and 4 are based on the following information:

A vehicle with its windows rolled up is traveling down the road at 50 miles an hour. Two balloons are inside. One balloon is hanging straight down from the ceiling by a string. The other balloon is also attached to a string but is floating straight up (see figure). When the driver slams on the brakes, the hanging balloon swings forward and the floating balloon swings backward.

This observation raises an interesting question: Why did the hanging balloon go forward while the floating balloon went backward? Here is a possible explanation: The hanging balloon is relatively heavy; so its momentum carried it forward when the vehicle stopped. The floating balloon, being lighter than air and having less momentum, went backward because as the vehicle stopped, the heavier air molecules inside the vehicle rushed forward and piled up at the front. Thus, the piled-up air molecules at the front pushed harder on the front side of the balloon than the relatively fewer air molecules on the balloon's backside. Thus, the balloon was pushed backward.

3. Suppose you have two balloons just like those in the vehicle, a large airtight chamber on wheels, and a vacuum pump (a pump that can extract air from airtight chambers). What experiment using these materials would test the possible explanation?

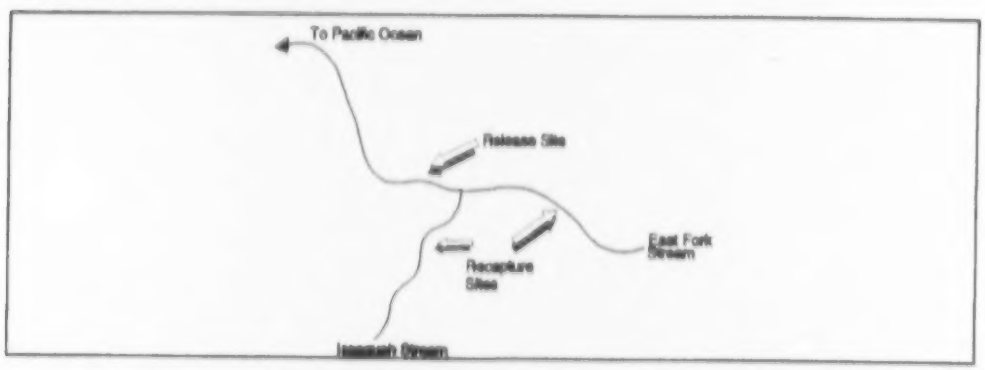
a. Suck the air out of the chamber. Because air does not weigh anything, nothing will happen to either balloon.
b. Attach the two balloons inside the chamber. Extract the air. Push the chamber and then stop it.
c. Replicate the experiment using the vehicle just as before so that you have a controlled experiment.
d. Place the balloons in the chamber and set it in motion. Then stop it and use the pump to extract the air.
e. The hanging balloon is heavier so it will swing with the momentum. The floating balloon is lighter so it falls back.

4. What result of your experiment would show that the explanation is probably wrong?

   a. The momentum will carry the heavier balloon forward.
   b. The two balloons each do something different.
   c. If both balloons moved in the direction of the vacuum at the same pace and stopped at the same time, then it does not matter if there are molecules pushing on the balloons.
   d. The balloons would go backward.
   e. The floating balloon goes backward.

Items 5 - 11 are based on the following information:

To test the hypothesis that salmon return to their home stream to spawn using their sense of smell, a biologist captured returning East Fork and Issaquah salmon at the two recapture points marked on the map. He then plugged the noses of some of the fish from both streams (the experimental fish) and he left the noses of the other fish unplugged (the control fish). He then took all the fish to the release point (as marked on the map). The fish then swam up stream and were recaptured in the East Fork or Issaquah streams at the two points marked recapture points. The biologist's data are shown in Tables 1 and 2.



Map of Issaquah and East Fork streams showing salmon release and recapture sites.

TABLE 1: RESULTS FOR EXPERIMENTAL FISH WITH PLUGGED NOSES

	Recapture Site	
Homestream	Issaquah	East Fork
Issaquah	39	12
East Fork	16	3

TABLE 2: RESULTS FOR CONTROL FISH WITH UNPLUGGED NOSES

	Recapture Site	
Homestream	Issaquah	East Fork
Issaquah	46	0
East Fork	8	19

5. Based on the results in Table 1, what percentage of fish made the wrong turn at the fork as they swam back up stream?

   a. 89%
   b. 46%
   c. 19%
   d. 11%
   e. 8%

6. Based on the results in Table 2, what percentage of fish made the wrong turn?

   a. 77%
   b. 60%
   c. 40%
   d. 28%
   e. 16%

7. Assuming that salmon navigate by smell, what percentage of the experimental fish would you expect to be recaptured in their home stream?

   a. 100%
   b. 60%
   c. 50%
   d. 25%
   e. 0%

8. Do the results in Tables 1 and 2 support the smell hypothesis?

   a. no, because some of control fish made the wrong turn and some of the experimental fish made the correct turn
   b. yes, because a significantly greater percent of experimental fish made the wrong turn than did control fish
   c. yes, because none of the fish with plugged noses found their home stream
   d. no, because a significantly greater percent of experimental fish did not make the wrong turn
   e. can not tell because the sample was too small

9. What is the independent variable in the experiment?

   a. the release point
   b. the recapture point
   c. fishes' ability to smell
   d. fishes' ability to return to home stream
   e. fishes' ability to see

10. What is the dependent variable in the experiment?

   a. the release point
   b. the recapture point
   c. fishes' ability to smell
   d. fishes' ability to see
   e. original place of capture

11. Which variable should be held constant in this experiment?

a. original place of capture
b. the recapture point
c. fishes' ability to smell
d. fishes' ability to return to home stream
e. fishes' ability to see

APPENDIX VI

NATURE OF SCIENCE IN NOVEL CONTEXTS

Next to each item write the number that best reflects your current belief:
1=strongly disagree 2=disagree 3=don't know 4=agree 5=strongly agree

1. Current scientific theories portray nature more accurately than those they replaced.

2. Current ideas about processes such as evolution, photosynthesis, and atomic structure will probably be modified in the future.

3. New explanations are accepted by scientists primarily on how well they "fit" with accepted explanations in related fields.

4. Scientists think that atoms exist primarily because they have seen them through powerful microscopes.

5. New discoveries depend mostly on luck.

6. Hypothesis formation involves creativity.

7. To conclude that a hypothesis has been "supported" or "not supported," one must first compare observations with expectations.

APPENDIX VII

STUDENT SATISFACTION SURVEY

Please answer the following questions on a scale ranging from 1 to 5 (or from A to E on the answer sheet), with 1 (A) corresponding to the most positive response, and 5 (E) the most negative response.

1. (ATTENTION) Does the course hold your attention and interest?

2. (LEVEL) Does the instructor cover the material at an intelligent level and a pace appropriate to you?

3. (IMPACT) Does the class stimulate your intellectual curiosity?

4. (CLARITY) Does the instructor present the material in a clear and understandable manner?

5. (EMPHASIS) Does the instructor make it clear to you what is expected of you in the course?

6. (CREDIBILITY) Do you have confidence in the instructor's knowledge of the subject?

7. (SENSITIVITY) Does the instructor seem to know when the class is having trouble understanding the material?

8. (HELPFULNESS) Is the instructor helpful outside class?

9. (RESPONSIVENESS) Does the instructor encourage student response?

10. (ENTHUSIASM) Does the instructor convey enthusiasm about the course?

11. (EXAMINATIONS) Do the examinations reflect the assigned material?

12. (WILLINGNESS) Would you willingly take the course from another instructor?

13. (EVALUATION) What overall grade would you give the instructor?